



OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/998,904

DATE: 06/11/2002
TIME: 15:06:35

Input Set : A:\119929-1037 SEQ ID LISTING.txt
Output Set: N:\CRF3\06112002\I998904.raw

P.3 m

3 <110> APPLICANT: The University of Texas Board of Regents
5 <120> TITLE OF INVENTION: Prediction of Disease-Causing Alleles from Sequence Context
7 <130> FILE REFERENCE: 119929-1037
9 <140> CURRENT APPLICATION NUMBER: 09/998,904
10 <141> CURRENT FILING DATE: 2002-06-04
12 <160> NUMBER OF SEQ ID NOS: 12
14 <170> SOFTWARE: PatentIn version 3.1
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 101
18 <212> TYPE: DNA
19 <213> ORGANISM: mammal
21 <220> FEATURE:
22 <221> NAME/KEY: misc_feature
23 <222> LOCATION: (51)..(51)
24 <223> OTHER INFORMATION: n = a or g
26 <400> SEQUENCE: 1
27 gctgggcca gctctacagc ttggtgatct gggggtgtac gctgctcctg ngctcaccca 60
29 tgctggtgtt ccggaccatg aaggagtaca gcgatgaggg c 101
32 <210> SEQ ID NO: 2
33 <211> LENGTH: 101
34 <212> TYPE: DNA
35 <213> ORGANISM: mammal
37 <220> FEATURE:
38 <221> NAME/KEY: misc_feature
39 <222> LOCATION: (51)..(51)
40 <223> OTHER INFORMATION: n = c or g
42 <400> SEQUENCE: 2
43 gctggtgttc cggaccatga aggagtacag cgatgagggc cacaacgtca ncgcttgtgt 60
45 catcagctac ccacccctca tctgggaagt gttcaccaac a 101
48 <210> SEQ ID NO: 3
49 <211> LENGTH: 101
50 <212> TYPE: DNA
51 <213> ORGANISM: mammal
53 <220> FEATURE:
54 <221> NAME/KEY: misc_feature
55 <222> LOCATION: (51)..(51)
56 <223> OTHER INFORMATION: n = a or g
58 <400> SEQUENCE: 3
59 gggccaagct ctacagcttg gtgatctggg ggtgtacgct gctcctgagc ncacccatgc 60
61 tgggtgttccg gaccatgaag gagtacagcg atgagggcca c 101
64 <210> SEQ ID NO: 4
65 <211> LENGTH: 101
66 <212> TYPE: DNA

**Does Not Comply
Corrected Diskette Needed**

RAW SEQUENCE LISTING

DATE: 06/11/2002

PATENT APPLICATION: US/09/998,904

TIME: 15:06:35

Input Set : A:\119929-1037 SEQ ID LISTING.txt

Output Set: N:\CRF3\06112002\I998904.raw

```

67 <213> ORGANISM: mammal
69 <220> FEATURE:
70 <221> NAME/KEY: misc_feature
71 <222> LOCATION: (51)..(51)
72 <223> OTHER INFORMATION: n = a or g
74 <400> SEQUENCE: 4
W 75 ctgcccttct gggccatcac catctccaac aacttcgact ggctctttgg ngagacgctc 60
77 tgccgcgtgg tgaatgccat tatctccatg aacctgtaca g 101
80 <210> SEQ ID NO: 5
81 <211> LENGTH: 101
82 <212> TYPE: DNA
83 <213> ORGANISM: mammal
85 <220> FEATURE:
86 <221> NAME/KEY: misc_feature
87 <222> LOCATION: (51)..(51)
88 <223> OTHER INFORMATION: n = c or t
W 90 <400> SEQUENCE: 5
91 cttctggggc atcaccatct ccaacaactt cgactggctc tttggggaga ngctctgccg 60
93 cgtggtgaat gccattatct ccatgaacct gtacagcagc a 101
96 <210> SEQ ID NO: 6
97 <211> LENGTH: 101
98 <212> TYPE: DNA
99 <213> ORGANISM: mammal
101 <220> FEATURE:
102 <221> NAME/KEY: misc_feature
103 <222> LOCATION: (51)..(51)
104 <223> OTHER INFORMATION: n = g or a
W 106 <400> SEQUENCE: 6
107 ggacaccggc caccctccgc gccaccacc ctcgctttct ccggttcct ntggcccagg 60
109 cgccgcgcgg acccggcagc tgtctgcgca cgccgagctc c 101
112 <210> SEQ ID NO: 7
113 <211> LENGTH: 101
114 <212> TYPE: DNA
115 <213> ORGANISM: mammal
117 <220> FEATURE:
118 <221> NAME/KEY: misc_feature
119 <222> LOCATION: (51)..(51)
120 <223> OTHER INFORMATION: n = g or a
W 122 <400> SEQUENCE: 7
123 ctgtctcagc agtggacagc gggcatgggt ctgctgatgg cgtcctcgt nctgctcctc 60
125 gtggcgggca atgtgctggt gatcgtggcc atcgccaaga c 101
128 <210> SEQ ID NO: 8
129 <211> LENGTH: 101
130 <212> TYPE: DNA
131 <213> ORGANISM: mammal
133 <220> FEATURE:
134 <221> NAME/KEY: misc_feature
135 <222> LOCATION: (51)..(51)
136 <223> OTHER INFORMATION: n = t or c

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/998,904

DATE: 06/11/2002
TIME: 15:06:35

Input Set : A:\119929-1037 SEQ ID LISTING.txt
Output Set: N:\CRF3\06112002\I998904.raw

```

138 <400> SEQUENCE: 8
W-OK 139 cgggcatggg tctgtgatg gcgctcatcg tgctgtcat cgtggcgggc natgtgtgg      60
141 tgatcgtggc catcgccaag acgcgcgggc tgcagacgt c      101
144 <210> SEQ ID NO: 9
145 <211> LENGTH: 101
146 <212> TYPE: DNA
147 <213> ORGANISM: mammal
149 <220> FEATURE:
150 <221> NAME/KEY: misc_feature
151 <222> LOCATION: (51)..(51)
152 <223> OTHER INFORMATION: n = a or g
154 <400> SEQUENCE: 9
OK-> 155 ctgtgtgatg cgtctcatgt gctgtctatc gtggcgggca atgtgtgtgt natcgtggcc      60
157 atcgccaaga cgccgcgggt gcagacgtc accaactct t      101
160 <210> SEQ ID NO: 10
161 <211> LENGTH: 101
162 <212> TYPE: DNA
163 <213> ORGANISM: mammal
165 <220> FEATURE:
166 <221> NAME/KEY: misc_feature
167 <222> LOCATION: (51)..(51)
168 <223> OTHER INFORMATION: n = t or c
170 <400> SEQUENCE: 10
OK-> 171 tcgccaagac gccgcggctg cagacgtca ccaacctctt catcatgtcc ntggccagcg      60
173 cgcacctgtt catggggctg ctggtggtgc cgttcggggc c      101
176 <210> SEQ ID NO: 11
177 <211> LENGTH: 101
178 <212> TYPE: DNA
179 <213> ORGANISM: mammal
181 <220> FEATURE:
182 <221> NAME/KEY: misc_feature
183 <222> LOCATION: (51)..(51)
184 <223> OTHER INFORMATION: n = a or g
186 <400> SEQUENCE: 11
OK-> 187 cgtggtgtgg ggcgcgtggg agtacggctc cttcttctgc gagctgtgga nctcagtgga      60
189 cgtgtgtgtc gtgacggcca gcatcgagac cctgtgtgtc a      101
192 <210> SEQ ID NO: 12
193 <211> LENGTH: 100
194 <212> TYPE: DNA
195 <213> ORGANISM: mammal
W--> 197 <220> FEATURE: < delete
198 <221> NAME/KEY: misc_feature
199 <222> LOCATION: (51)..(51)
200 <223> OTHER INFORMATION: n = t or c
202 <400> SEQUENCE: 12
203 ttccgtacc agagcctgt gacgcgcgcg cgggcgcggg gcctcgtgtg accgtgtggg      60
205 ccattctggc cctggtgtcc ttctgcca tctcatgca      100

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/998,904

DATE: 06/11/2002
TIME: 15:06:36

Input Set : A:\119929-1037 SEQ ID LISTING.txt
Output Set: N:\CRF3\06112002\I998904.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 51
Seq#:2; N Pos. 51
Seq#:3; N Pos. 51
Seq#:4; N Pos. 51
Seq#:5; N Pos. 51
Seq#:6; N Pos. 51
Seq#:7; N Pos. 51
Seq#:8; N Pos. 51
Seq#:9; N Pos. 51
Seq#:10; N Pos. 51
Seq#:11; N Pos. 51